REMARKS

In view of the following discussion, the Applicants submit that none of the claims now pending in the application is anticipated under the provisions of 35 U.S.C. § 102 or made obvious under the provisions of 35 U.S.C. § 103. Thus, the Applicants believe that all of these claims are now in allowable form.

I. REJECTION OF CLAIMS 1-7, 11, 14-15, 32-38, 40 AND 43-45 UNDER 35 U.S.C. § 102

The Examiner has rejected claims 1-7, 11, 14-15, 32-38, 40 and 43-45 under 35 U.S.C. §102(e) as being anticipated by the Haskell patent (US patent 6,233,356, issued on May 15, 2001, hereinafter Haskell). In response, the Applicants have amended independent claims 1 and 32, from which claims 2-7, 11, 14-15, 33-38, 40 and 43-45 depend, to more clearly recite aspects of the present invention.

Haskell teaches a video coding system that produces a single embedded data stream containing several video object layers. In particular, Haskell teaches that an original video sequence is encoded into a data stream comprising a first or "base" layer containing coded video object data of a lowest quality (*e.g.*, temporal or spatial quality), and one or more "enhancement" layers containing enhancement data that, when combined with the base layer video, display increasingly higher-quality video sequences. These multiple video object layers are then organized into a single data stream by a multiplexer, and the data stream is sent to a decoder for display (*see*, Haskell, column 4, lines 37-39: "The MUX 600 organizes the coded video object data ... into a data stream ..."). The decoder extracts one or more of the video object data layers to display, in accordance with the decoder's own operating parameters. However, Haskell does not teach that coded video object layers are output to the decoder as a plurality of data streams.

The Examiner's attention is directed to the fact that Haskell fails to disclose or suggest the novel invention of separating a video image sequence into two or more components and then encoding each component of the video image sequence in

accordance with selected dimensions "to form a plurality of bitstreams", as claimed in Applicants' independent claims 1 and 32. Specifically, Applicants' claims 1 and 32, as amended, positively recite:

> 1. A method of deconstructing video comprising:
> separating a video image sequence into two or more components;
> selecting a plurality of dimensions, where each dimension represents a characteristic of the video image sequence; and
> encoding each component of the video image sequence in accordance with the selected dimensions to form a plurality of bitstreams. (Emphasis added)

> 32. A computer readable medium containing software that, when executed by one or more general purpose computers operating as network nodes, causes the computer or computers to perform a method comprising:
> separating a video image sequence into two or more components;
> selecting a plurality of dimensions, where each dimension represents a characteristic of the video image sequence; and
> encoding each component of the video image sequence in accordance with the selected dimensions to form a plurality of bitstreams. (Emphasis added)

Applicants' invention is directed to a method and apparatus for generating, distributing and reconstructing deconstructed video over a network. In contrast, conventional systems that distribute deconstructed video through communications networks often encode video sequences into several independent data streams, wherein each data stream represents an entire video sequence having a different level of image quality. An end user or decoder device then selects one appropriate data stream to match user equipment capabilities. Such methods typically require a priori user familiarity with device capabilities in order to manually select the appropriate data stream. Moreover, the transmission of multiple data streams containing full representations of video sequences consumes a tremendous amount of bandwidth and storage space.

The present invention provides a method for encoding video sequences in which a video sequence is divided into two or more constituent components (*e.g.*, foreground/background, moving objects/stationary objects, text/versus moving video,

face/remaining video, fixed regions/other regions, infrared geometric shapes/other regions, annotated regions/other regions, graphics/non-graphics, etc.). The components of the video sequence are then deconstructed into multiple dimensions (*e.g.*, resolution, frame rate, display type, etc) by generating a partial order representation of the deconstructed elements for each component of the video. The method thereby produces a common base bitstream calculated from common components of the individual dimensional base bitstreams, plus a plurality of additional or augmentation bitstreams, which collectively form a lattice structure. Intersection points of the bitstreams represent "improved" video sequences including combinations of the base bitstream with one or more of the augmentation bitstreams. Thus, by performing multidimensional video deconstruction, any two or more subsets of bitstreams may be combined to produce an optimal video sequence for a particular user device.

In contrast, Haskell only teaches a video encoder that performs scalable compression on an original image to produce a <u>single</u> embedded data stream from which a decoder may extract video data layers having different video qualities (*e.g.*, different temporal or spatial resolutions). In other words, every image frame in every video object layer in the output data stream is subjected to the same encoding parameter.

In contrast, the Applicants' invention positively claims the step of encoding <u>each component</u> (of two or more components) of the video image sequence (in accordance with selected dimensions or encoding parameters) <u>to form a plurality of bitstreams</u>. This allows, for example, two or more components of the video sequence to be deconstructed <u>according to different dimensions</u> (*e.g.*, the background may be deconstructed according to frame rate, and the foreground may be deconstructed according to resolution and so on). In addition, each component may be deconstructed according to a plurality of dimensions (*e.g.*, each dimension represented by a separate bitstream), and a combination of bitstreams may be selected that allows a video sequence having specified qualities in a <u>plurality</u> of dimensions to be reconstructed. Haskell's system is completely devoid of any teaching relating to the need or desire to

encode each component of a video sequence into a plurality of deconstructed bitstreams.

Therefore, the Applicants submit that, at least for the reasons presented above, independent claims 1 and 32, as amended, fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

Dependent claims 2-7, 11, 14-15, 33-38, 40 and 43-45 depend from claims 1 and 32, and recite additional features therefore. As such, and for at least the same reasons set forth above, the Applicants submit that claims 2-7, 11, 14-15, 33-38, 40 and 43-45 are not anticipated by the teachings of Haskell. Therefore, the Applicants submit that dependent claims 2-7, 11, 14-15, 33-38, 40 and 43-45 also fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

II. REJECTION OF CLAIMS 8, 9 AND 25-28 UNDER 35 U.S.C. § 103

A. Claims 8 and 9

The Examiner rejected claims 8 and 9 under 35 U.S.C. §103(a) as being unpatentable over Haskell in view the Chaddha patent (U.S. Patent No. 5,621,660, issued April 15, 1997, hereinafter Chaddha). In response, the Applicants have amended independent claim 1, from which claims 8 and 9 depend, as discussed above to more clearly recite aspects of the invention.

Haskell has been discussed above.

Chaddha teaches a video delivery system that provides end-to-end encoding such that a single embedded data stream is produced containing several video sequence layers. In particular, Chaddha teaches that an original video sequence is encoded into a data stream comprising a first or "base" layer containing a video sequence of a lowest resolution, and two "enhancement" layers containing enhancement data that, when combined with the base layer video, display increasingly higher-resolution video sequences. A decoder may then extract one or more of the video sequence layers to display, in accordance with the decoder's own operating parameters. However, Chaddha, like Haskell, does <u>not</u> teach that two or more components of coded video sequence layers are each output to the decoder as a

plurality of data streams.

The Examiner's attention is directed to the fact that Chaddha, like Haskell, fails to disclose or suggest the novel invention of separating a video image sequence into two or more components and then encoding each component of the video image sequence in accordance with selected dimensions "to form a plurality of bitstreams", as claimed in Applicants' independent claim 1, which has been recited above.

As discussed above, the Applicants' invention provides a method for encoding video sequences in which a video sequence is divided into two or more constituent components which are then deconstructed into multiple dimensions by generating a partial order representation of the deconstructed elements for each component of the video. The method thereby produces a common base bitstream calculated from common components of the individual dimensional base bitstreams, plus a plurality of additional or augmentation bitstreams. By performing multidimensional video deconstruction, any two or more subsets of bitstreams may be combined to produce an optimal video sequence for a particular user device.

In contrast, both Haskell and Chaddha only teach a video encoder that performs scalable compression on an original image to produce a single embedded data stream from which a decoder may extract video data layers having different video qualities (*e.g.*, different temporal or spatial resolutions). In other words, every image frame in every video object layer in the output data stream is subjected to the same encoding parameter.

In contrast, the Applicants' invention positively claims the step of encoding each component (of two or more components) of the video image sequence (in accordance with selected dimensions or encoding parameters) to form a plurality of bitstreams. As discussed above, this allows two or more components of the video sequence to be deconstructed according to different dimensions, or allows each component to be deconstructed according to a plurality of dimensions. Thus, a combination of bitstreams may be selected that allows a video sequence having specified qualities in a plurality of dimensions to be reconstructed. Both Haskell's and Chaddha's systems are completely devoid of any teaching relating to the need or desire to encode each of at least two

components of a video sequence into a plurality of deconstructed bitstreams.

Therefore, as discussed above, the Applicants submit that, and at least for the reasons presented above, amended independent claim 1 fully satisfies the requirements of 35 U.S.C. §103 and is patentable thereunder.

Dependent claims 8 and 9 depend from claim 1, and recite additional features therefore. As such, and for at least same reasons set forth above, the Applicants submit that claims 8 and 9 are not made obvious by the teachings of Haskell in view of Chaddha. Therefore, the Applicants submit that dependent claims 8 and 9 also fully satisfy the requirements of 35 U.S.C. §103 and are patentable thereunder.

B. Claims 25-27

The Examiner rejected claims 25-27 under 35 U.S.C. §103(a) as being unpatentable over Chaddha in view the Burt patent (U.S. Patent No. 5,063,603, issued November 5, 1991, hereinafter Burt). In response, the Applicants have amended independent claim 25, from which claims 26-27 depend, to more clearly recite aspects of the invention.

Chaddha has been discussed above.

Burt teaches a method for object recognition, *e.g.*, for recognizing or locating an individual within a series of video frames. For example, a time series of successive, relatively high-resolution image frames, is examined in order to recognize the identity of a specific individual or object in the time series. The frames may then be examined at various resolutions to detect whether any earlier occurring frames include a group of attributes or image features possessed by an image of the specific individual or object. The locations of detected image features are then stored and used in subsequent higher-resolution frames to direct examination only to the image region of the detected features, *e.g.*, in order to verify the detection of the image features and/or to detect additional features or attributes of the image of the specific individual or object. Locations of stationary objects may also be stored in order to distinguish moving objects or images. By repeating this process for successive frames, the accumulated detected features can be used to recognize the detected image region as an image of the

specific individual or object.

The Examiner's attention is directed to the fact that Chaddha and Burt (either singly or in any permissible combination) fail to disclose or a method of separating a video image sequence into two or more components and then encoding each component of the video image sequence in accordance with a plurality of selected dimensions to form a plurality of bitstreams, as positively claimed by the Applicants. Applicants' independent claim 25 positively recites:

> 25. Apparatus for producing deconstructed video comprising:
> a video component extractor for extracting at least one second image sequence from a first image sequence, where said at least one second image sequence represents a component of said first video image sequence;
> an encoding dimension selector for selecting a plurality of dimensions to use to encode said at least one second image sequence; and
> a dimension-based encoder, coupled to said encoding dimension selector, for encoding the at least one second video image sequence into a plurality of bitstreams. (Emphasis added)

As recited in the preceding claim, Applicants' invention teaches an apparatus for producing deconstructed video that is capable of encoding components of a video image sequence in accordance with a plurality of dimensions. This approach allows various components of the video image sequence to be encoded in accordance with different dimensions, or allows any single component to be encoded in accordance with multiple dimensions, thereby optimizing a content match to an end user device (*e.g.*, decoder).

In contrast, neither Chaddha nor Burt teaches or suggests this novel approach. Both Chaddha and Burt teach deconstructing a video image sequence in accordance with only a single, particular dimension, namely, resolution. Burt is capable of identifying stationary objects as cited by the Examiner, but there is no teaching that each stationary object will be encoded differently than non-stationary objects. In other words, neither Chaddha nor Burt teaches or suggests that multidimensional deconstruction can be applied to components of the input image sequence. Combining the two references simply would not make Applicants' invention obvious.

Moreover, there is no suggestion or motivation to combine Chaddha and Burt in a manner that would yield the claimed invention. Chaddha is directed toward a method of distributing video over a network to a plurality of end users with varying decoding capabilities. Burt is directed toward a method of locating and/or identifying a particular individual or object within a series of successive video frames. It does not follow that an individual seeking to improve upon the video distribution system of Chaddha would look to the object recognition method of Burt. Likewise, an individual seeking to improve upon a method for identifying objects in a sequence of video images would not look to a video delivery system. Thus, the Applicants respectfully submit that the Examiner is clearly using hindsight to pick and choose elements from the references to support his rejection.

It is impermissible to use the claims as a framework from which to choose among individual references to recreate the claimed invention. *W. L. Gore Associates, Inc. v. Garlock, Inc.*, 220 U.S.P.Q. 303, 312 (1983). Moreover, the mere fact that a prior art structure could be modified to produce the claimed invention would not have made the modification obvious unless the <u>prior art suggested the desirability of the modification</u>. *In re Fritch*, 23 U.S.P.Q. 2d 1780, 1783, Fed. Cir. (1992); *In re Gordon*, 221 U.S.P.Q. 1125, 1127, Fed. Cir. (1984) (emphasis added). The rules applicable for combining references provide that there must be a suggestion from <u>within the references</u> to make the combination. *Uniroyal v. Rudkin-Wiley*, 5 U.S.P.Q. 2d 1434, 1438 (Fed. Cir. 1988); *In re Fine*, 5 U.S.P.Q. 2d at 1599 (emphasis added). Therefore, the teachings of Chaddha do not provide any justification for combination with the object recognition methodology of Burt. Thus, at least for the reasons presented above, independent claim 25 is not made obvious by the teachings of Chaddha in view of Burt.

Dependent claims 26 and 27 depend, either directly or indirectly, from claim 25, and recite additional features thereof. As such and for at least the same reasons set forth above, the Applicants submit that claims 26 and 27 are also not made obvious by the teachings of Chaddha in view of Burt. Therefore, the Applicants submit that all these dependent claims also fully satisfy the requirements of 35 U.S.C. § 103 and are patentable thereunder.

C. Claim 28

The Examiner rejected claim 28 under 35 U.S.C. §103(a) as being unpatentable over Chaddha and Burt in view of the Wine patent (U.S. Patent No. 6,477,201, issued November 5, 2002, hereinafter Wine). In response, the Applicants have amended independent claim 25, from which claim 28 depends, as discussed above to more clearly recite aspects of the invention.

Chaddha and Burt have been discussed above.

Wine teaches a method for selective enhancement or degradation of information within a video image sequence. For example, regions of particular interest within an image may be encoded with a higher resolution than regions of less significance, in order to emphasize the regions of interest. Thus, a single video image sequence is produced in which certain aspects of the image sequence have varying characteristic qualities. This is in contrast to a system which produces a plurality of quality-varying versions of the same video image sequence, wherein the quality is substantially uniform within a particular image sequence (*e.g.*, as taught by Chaddha).

Again, the gap left by Chaddha and Burt is not bridged by Wine. Namely, Wine also does not teach separating components of a video image sequence and encoding each of these components in accordance with a plurality of dimensions. Moreover, there is no suggestion or motivation to combine Chaddha and Burt with Wine in a manner that would yield the claimed invention. As discussed, the method taught by Wine teaches the creation of a single video image sequence having portions of varying quality within the same image sequence. In other words, there is no way to select different quality dimensions from the single image sequence for different components. Chaddha teaches the creation of a plurality of quality-varying video image sequences and enhancement files, wherein each sequence or enhancement file is characterized by a substantially uniform degree of quality. It does not follow that an individual seeking to improve upon the video distribution system of Chaddha would look to the selective encoding method of Wine. Neither does Chaddha teach or suggest that a method such as that taught by Wine could be incorporated in a way that would be advantageous or desirable. Likewise, an individual seeking to improve upon a method for emphasizing or

de-emphasizing objects in a sequence of video images would not look to a video delivery system that produces a plurality of quality-varying image sequences. Thus, the Applicants respectfully submit that the Examiner is clearly using hindsight to pick and choose elements from the references to support his rejection. Thus, for at least for the reasons presented above, independent claims 25 is not made obvious by the teachings of Chaddha in view of Burt and further in view of Wine.

Dependent claim 28 depends from claim 25 and recites additional features therefor. As such and for at least the same reasons set forth above, the Applicants submit that claim 28 is also not made obvious by the teachings of Chaddha in view of Burt and further in view of Wine. Therefore, the Applicants submit that claim 28 also fully satisfies the requirements of 35 U.S.C. § 103 and is patentable thereunder.

III. VOLUNTARY AMENDMENTS

The Applicants have voluntarily amended claims 6, 37 and 43 to address minor typographical errors.

IV. CONCLUSION

Thus, the Applicants submit that all of these claims now fully satisfy the requirements of 35 U.S.C. §102 and §103. Consequently, the Applicants believe that all these claims are presently in condition for allowance. Accordingly, both reconsideration of this application and its swift passage to issue are earnestly solicited.

If, however, the Examiner believes that there are any unresolved issues requiring the maintenance of the present final action in any of the claims now pending in the application, it is requested that the Examiner telephone Mr. Kin-Wah Tong, Esq. at (732) 530-9404 so that appropriate arrangements can be made for resolving such issues as expeditiously as possible.

09/707,044

Respectfully submitted,

9/20/04
Date

Kin-Wah Tong, Attorney
Reg. No. 39,400
(732) 530-9404

Moser, Patterson & Sheridan, LLP
595 Shrewsbury Avenue
Shrewsbury, New Jersey 07702